SEQUENCE LISTING

<110> CAMPBELL, ANTHONY KEITH

<120> PROTEIN AND DNA CODING THEREFOR

<130> WCM.69.US

<140> 09/831,142
<141> 2001-05-07

<150> PCT/GB99/03654
<151> 1999-11-05

<150> GB 9824357.9
<151> 1998-11-07

<160> 42

<170> PatentIn Ver. 2.1

<210> 1
<211> 870
<212> DNA
<213> Pholas dactylus

<220>
<221> CDS
<222> (30)..(704)

<400> 1

```
gaattcggca cgagtcggaa aagaacaaa atg gct tgt atc gtt ttc gtt gct   53
                                Met Ala Cys Ile Val Phe Val Ala
                                  1               5

ctt gtc gct cta tgc tta atg caa ccg ggt tcc ggt gag gaa gta caa   101
Leu Val Ala Leu Cys Leu Met Gln Pro Gly Ser Gly Glu Glu Val Gln
     10                  15                  20

tgc gcg atg aat tgg aca caa gct aat gaa tat gtg ttc aac gtg gac   149
Cys Ala Met Asn Trp Thr Gln Ala Asn Glu Tyr Val Phe Asn Val Asp
 25                  30                  35                  40

tgg atg acc att ttc atc tac gac tat ggc gct caa gag caa ctg tac   197
Trp Met Thr Ile Phe Ile Tyr Asp Tyr Gly Ala Gln Glu Gln Leu Tyr
                 45                  50                  55

gaa gat cgg gct ttg ggg ctg tgt cgg att gaa cgg gcc ggc cca ggt   245
Glu Asp Arg Ala Leu Gly Leu Cys Arg Ile Glu Arg Ala Gly Pro Gly
             60                  65                  70

acc aca aaa gcc gtc tgg att aac tgg agt aac gac acg cag tca tgt   293
Thr Thr Lys Ala Val Trp Ile Asn Trp Ser Asn Asp Thr Gln Ser Cys
         75                  80                  85

gta aca aga aaa aca atc ttc ttc gag gtt ggt gga gaa att gcc cgg   341
Val Thr Arg Lys Thr Ile Phe Phe Glu Val Gly Gly Glu Ile Ala Arg
     90                  95                  100
```

```
cta gtt gac tac aga cca cag gaa gac gga act gag aaa act ttt aca   389
Leu Val Asp Tyr Arg Pro Gln Glu Asp Gly Thr Glu Lys Thr Phe Thr
105                 110                 115                 120

aga aaa ttc tct agc aaa atg cca ggc act tac atg ctt atg gac gtg   437
Arg Lys Phe Ser Ser Lys Met Pro Gly Thr Tyr Met Leu Met Asp Val
                125                 130                 135

tgc gct aca agg gac gct gat gat aaa tgc atc gaa ggc aca att gtg   485
Cys Ala Thr Arg Asp Ala Asp Asp Lys Cys Ile Glu Gly Thr Ile Val
            140                 145                 150

gtg aca gtc agg gtg tcc cta tat gac gaa gat aac aat ggt gta atg   533
Val Thr Val Arg Val Ser Leu Tyr Asp Glu Asp Asn Asn Gly Val Met
        155                 160                 165

gat gaa ggt aag gtg att cca tct gag aca atc gag gat gat atc aag   581
Asp Glu Gly Lys Val Ile Pro Ser Glu Thr Ile Glu Asp Asp Ile Lys
    170                 175                 180

gac tgt ggg ctc tta gac caa gat gtt gaa ctc gat tat acg tgg act   629
Asp Cys Gly Leu Leu Asp Gln Asp Val Glu Leu Asp Tyr Thr Trp Thr
185                 190                 195                 200

caa aac gag tgt gat cta cca gac aca gta gac gag gct gaa gac aca   677
Gln Asn Glu Cys Asp Leu Pro Asp Thr Val Asp Glu Ala Glu Asp Thr
                205                 210                 215

ccg tca gaa act gga gaa ttc ttc tgg tagatctatc agactacttt         724
Pro Ser Glu Thr Gly Glu Phe Phe Trp
            220                 225

tatcagcagg acaactggtc gttaccagac acctataacg tgtcctcatc aataatgtgt 784

aaaacagaaa taatcgatag aatattgaaa ataaaatgtt aataaacact ggttgaaata 844

tgaaaaaaaa aaaaaaaaaa ctcgag                                      870
```

<210> 2
<211> 816
<212> DNA
<213> Pholas dactylus

<400> 2

```
gaattcggca cgagggaaaa gaacaaaatg gcttgtatcg ttttcgttgc tcttgtcgct 60
ctatgcttaa tgcaaccggg ttccggtgag gaagtacaat gcgcgatgaa ttggacacaa 120
gctaatgaat atgtgttcaa cgtggactgg atgaccattt tcatctacga ctatggcgct 180
caagagcaac tgtacgagga tcgggctttg gggctgtgtc ggattgaacg ggccggccca 240
ggtaccacaa aagccgtctg gattaactgg agtaacgaca cgcagtcatg tgtaacaaga 300
aaaacaatct tcttcgaggt tggtggagaa attgcccggc tagttgacta cagaccacag 360
gaagacggaa ctgagaaaac ttttacaaga aaattctcta gcaaaatgcc aggcacttac 420
atgcttatgg acgtgtgcgc tacaagggac gctgatgata aatgcatcga aggcacaatt 480
gtggtgacag tcagggtgtc cctatatgac gaagataaca atggtgtaat ggatgaaggt 540
aaggttattc catctgagac aatcgaggat gatatcaagg actgtgggct cttagaccaa 600
gatgttgaac tcgattatac gtggactcaa aacgagtgtg atctaccaga cacagtagac 660
gaggctgaag acacaccgtc agaaactgga gaattcttct ggtagatcta tcagaccact 720
tttatcagca ggacaactgg tcgttaccag acacctataa cgtgtcctca tcaataatgt 780
```

gtaaaacaga aataatcgat agaatattga aaataa 816

<210> 3
<211> 852
<212> DNA
<213> Pholas dactylus

<400> 3
gtcggaaaag aacaaaatgg cttgtatcgt tttcgttgct cttgtcgctc tatgcttaat 60
gcaaccgggt tccggtgagg aagtacaatg cgcgatgaat tggacacaag ctaatgaata 120
tgtgttcaac gtggactgga tgaccatttt catctacgac tatggcgctc aagagcaact 180
gtacgaggat cgggctttgg ggctgtgtcg gattgaacgg gccggcccag gtaccacaaa 240
agccgtctgg attaactgga gtaacgacac gcagtcatgt gtaacaagaa aaacaatctt 300
cttcgaggtt ggtggagaaa ttgcccggct agttgactac agaccacagg aagacggaac 360
tgagaaaact tttacaagaa aattctctag caaaatgcca ggcacttaca tgcttatgga 420
cgtgtgcgct acaagggacg ctgatgataa atgcatcgaa ggcacaattg tggtgacagt 480
cagggtgtcc ctatatgacg aagataacaa tggtgtaatg gatgaaggta aggttattcc 540
atctgagaca atcgaggatg atatcaagga ctgtgggctc ttagaccaag atgttgaact 600
cgattatacg tggactcaaa acgagtgtga tctaccagac acagtagacg aggctgaaga 660
cacaccgtca gaaactggag aattcttctg gtagatctat cagaccactt ttatcagcag 720
gacaactggt cgttaccaga cacctataac gtgtcctcat caataatgtg taaaacagaa 780
ataatcgata gaatattgaa aataaaatgt taatagacac tggttgaaaa aaaaaaaaaa 840
aaaaaactcg ag 852

<210> 4
<211> 225
<212> PRT
<213> Pholas dactylus

<400> 4
Met Ala Cys Ile Val Phe Val Ala Leu Val Ala Leu Cys Leu Met Gln
1 5 10 15

Pro Gly Ser Gly Glu Glu Val Gln Cys Ala Met Asn Trp Thr Gln Ala
20 25 30

Asn Glu Tyr Val Phe Asn Val Asp Trp Met Thr Ile Phe Ile Tyr Asp
35 40 45

Tyr Gly Ala Gln Glu Gln Leu Tyr Glu Asp Arg Ala Leu Gly Leu Cys
50 55 60

Arg Ile Glu Arg Ala Gly Pro Gly Thr Thr Lys Ala Val Trp Ile Asn
65 70 75 80

Trp Ser Asn Asp Thr Gln Ser Cys Val Thr Arg Lys Thr Ile Phe Phe
85 90 95

Glu Val Gly Gly Glu Ile Ala Arg Leu Val Asp Tyr Arg Pro Gln Glu
100 105 110

Asp Gly Thr Glu Lys Thr Phe Thr Arg Lys Phe Ser Ser Lys Met Pro
115 120 125

Gly Thr Tyr Met Leu Met Asp Val Cys Ala Thr Arg Asp Ala Asp Asp
130 135 140

```
Lys Cys Ile Glu Gly Thr Ile Val Val Thr Val Arg Val Ser Leu Tyr
145                 150                 155                 160

Asp Glu Asp Asn Asn Gly Val Met Asp Glu Gly Lys Val Ile Pro Ser
                165                 170                 175

Glu Thr Ile Glu Asp Asp Ile Lys Asp Cys Gly Leu Leu Asp Gln Asp
            180                 185                 190

Val Glu Leu Asp Tyr Thr Trp Thr Gln Asn Glu Cys Asp Leu Pro Asp
        195                 200                 205

Thr Val Asp Glu Ala Glu Asp Thr Pro Ser Glu Thr Gly Glu Phe Phe
    210                 215                 220

Trp
225
```

<210> 5
<211> 205
<212> PRT
<213> Pholas dactylus

<400> 5

```
Glu Glu Val Gln Cys Ala Met Asn Trp Thr Gln Ala Asn Glu Tyr Val
  1               5                  10                  15

Phe Asn Val Asp Trp Met Thr Ile Phe Ile Tyr Asp Tyr Gly Ala Gln
            20                  25                  30

Glu Gln Leu Tyr Glu Asp Arg Ala Leu Gly Leu Cys Arg Ile Glu Arg
        35                  40                  45

Ala Gly Pro Gly Thr Thr Lys Ala Val Trp Ile Asn Trp Ser Asn Asp
    50                  55                  60

Thr Gln Ser Cys Val Thr Arg Lys Thr Ile Phe Phe Glu Val Gly Gly
65                  70                  75                  80

Glu Ile Ala Arg Leu Val Asp Tyr Arg Pro Gln Glu Asp Gly Thr Glu
                85                  90                  95

Lys Thr Phe Thr Arg Lys Phe Ser Ser Lys Met Pro Gly Thr Tyr Met
            100                 105                 110

Leu Met Asp Val Cys Ala Thr Arg Asp Ala Asp Asp Lys Cys Ile Glu
        115                 120                 125

Gly Thr Ile Val Val Thr Val Arg Val Ser Leu Tyr Asp Glu Asp Asn
    130                 135                 140

Asn Gly Val Met Asp Glu Gly Lys Val Ile Pro Ser Glu Thr Ile Glu
145                 150                 155                 160

Asp Asp Ile Lys Asp Cys Gly Leu Leu Asp Gln Asp Val Glu Leu Asp
                165                 170                 175
```

```
Tyr Thr Trp Thr Gln Asn Glu Cys Asp Leu Pro Asp Thr Val Asp Glu
            180                 185                 190

Ala Glu Asp Thr Pro Ser Glu Thr Gly Glu Phe Phe Trp
        195                 200                 205


<210> 6
<211> 225
<212> PRT
<213> Pholas dactylus

<400> 6
Met Ala Cys Ile Val Phe Val Ala Leu Val Ala Leu Cys Leu Met Gln
  1               5                  10                  15

Pro Gly Ser Gly Glu Glu Val Gln Cys Ala Met Asn Trp Thr Gln Ala
             20                  25                  30

Asn Glu Tyr Val Phe Asn Val Asp Trp Met Thr Ile Phe Ile Tyr Asp
         35                  40                  45

Tyr Gly Ala Gln Glu Gln Leu Tyr Glu Asp Arg Ala Leu Gly Leu Cys
     50                  55                  60

Arg Ile Glu Arg Ala Gly Pro Gly Thr Thr Lys Ala Val Trp Ile Asn
 65                  70                  75                  80

Trp Ser Asn Asp Thr Gln Ser Cys Val Thr Arg Lys Thr Ile Phe Phe
                 85                  90                  95

Glu Val Gly Gly Glu Ile Ala Arg Leu Val Asp Tyr Arg Pro Gln Glu
            100                 105                 110

Asp Gly Thr Glu Lys Thr Phe Thr Arg Lys Phe Ser Ser Lys Met Pro
        115                 120                 125

Gly Thr Tyr Met Leu Met Asp Val Cys Ala Thr Arg Asp Ala Asp Asp
    130                 135                 140

Lys Cys Ile Glu Gly Thr Ile Val Val Thr Val Arg Val Ser Leu Tyr
145                 150                 155                 160

Asp Glu Asp Asn Asn Gly Val Met Asp Glu Gly Lys Val Ile Pro Ser
                165                 170                 175

Glu Thr Ile Glu Asp Asp Ile Lys Asp Cys Gly Leu Leu Asp Gln Asp
            180                 185                 190

Val Glu Leu Asp Tyr Thr Trp Thr Gln Asn Glu Cys Asp Leu Pro Asp
        195                 200                 205

Thr Val Asp Glu Ala Glu Asp Thr Pro Ser Glu Thr Gly Glu Phe Phe
    210                 215                 220

Trp
225
```

<210> 7
<211> 17
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<220>
<221> modified_base
<222> (3)
<223> i

<400> 7
acnathttyt tycargt 17

<210> 8
<211> 17
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<220>
<221> modified_base
<222> (12)
<223> A, T, C or G

<220>
<221> modified_base
<222> (15)
<223> i

<400> 8
cargargarg gnacnga 17

<210> 9
<211> 17
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<220>
<221> modified_base
<222> (3)
<223> i

<220>
<221> modified_base
<222> (6)
<223> A, T, C or G

<400> 9
tcngtnccyt cytcytg 17

<210> 10
<211> 18
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<220>
<221> modified_base
<222> (9)
<223> i

<400> 10
ttyaaygtng aytggatg 18

<210> 11
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 11
acacagcccc aaagcccgat 20

<210> 12
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 12
ttgcccggct agttgactac 20

<210> 13
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 13
catatttcaa ccagtgttta ttaa 24

<210> 14
<211> 19
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 14
caattgtgcc ttcgatgca 19

<210> 15
<211> 17
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 15
ggactgtggg ctcttag 17

<210> 16
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 16
atggcttgta tcgttttcgt 20

<210> 17
<211> 27
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 17
ccacacggat cctgaggaag tacaatg 27

<210> 18
<211> 27
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 18
ccacacggat ccttattgat gaggaca 27

<210> 19
<211> 53
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 19
cttgttttta tggtcgtcta catttcttac atctatgcgg aggaagtaca atg 53

<210> 20
<211> 54
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 20
ccacacagat ctagaatgaa attcttagtc aacgttgccc ttgtttttat ggtc 54

<210> 21
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 21
tttactgttt tcgtaacagt tttg 24

<210> 22
<211> 20

<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 22
caacaacgca cagaatctag 20

<210> 23
<211> 726
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Clone 40 amplified by rTth DNA pol XL

<220>
<221> modified_base
<222> (644)
<223> A, T, C, G, other or unknown

<400> 23
tatgcttaat gcaaccgggt tccggtgagg aagtacaatg cgcgatgaat tggacacaag 60
ctaatgaata tgtgttcaac gtggactgga tgaccatttt catctacgac tatggcgctc 120
aagagcaact gtacgaagat cgggctttgg ggctgtgtcg gattgaacgg gccggcccag 180
gtaccacaaa agccgtctgg attaactgga gtaacgacac gcagtcatgt gtaacaagaa 240
aaacaatctt cttcgaggtt ggtggagaaa ttgcccggct agttgactac agaccacagg 300
aagacggaac tgagaaaact tttacaagaa aattctctag caaaatgcca ggcacttaca 360
tgcttatgga cgtgtgcgct acaagggacg ctgatgataa atgcatcgaa ggcacaattg 420
tggtgacagt cagggtgtcc ctatatgacg aagataacaa tggtgtaatg gatgaaggta 480
aggtgattcc atctgagaca atcgaggatg atatcaagga ctgtgggctc ttagaccaag 540
atgttgaact cgattatacg tggactcaaa acgagtgtga tctaccagac acagtagacg 600
aggctgaaga cacaccgtca gaaactggag aattcttctg gtanatctat cagactactt 660
ttatcagcag gacaactggt cgttaccaga cacctataac gtgtcctcat caataatgtg 720
taaaac 726

<210> 24
<211> 34
<212> PRT
<213> Saccharomyces cerevisiae

<400> 24
Asn Leu Arg Asp Glu Asp Asn Asn Leu Leu Asp Glu Asn Gly Asp Leu
1 5 10 15

Leu Pro Leu Glu Ser Leu Glu Leu Asp Gln Asp Val Glu Leu Asp Tyr
20 25 30

Thr Trp

<210> 25
<211> 31

<212> PRT
<213> Cyprinus carpio

<400> 25
Ile Met Gln Lys Gly Glu Leu Val Pro Leu Asp Thr Val Leu Asp Met
1 5 10 15

Ile Lys Asp Ala Met Ile Ala Lys Ala Asp Val Ser Lys Gly Tyr
20 25 30

<210> 26
<211> 20
<212> PRT
<213> Synechocystis sp.

<400> 26
Asp Gln Val Gln Ser Leu Met Arg Phe Ser Gln Ser Lys Gln Ile Ile
1 5 10 15

Phe Asn Phe Asp
20

<210> 27
<211> 14
<212> PRT
<213> Emericella nidulans

<400> 27
Ile Met Cys Ser Val Asp Trp Thr Arg Arg Asn Arg Phe Ile
1 5 10

<210> 28
<211> 14
<212> PRT
<213> Drosophila melanogaster

<400> 28
Pro Asp Thr Tyr Asp Glu Glu Glu Asp Thr Tyr Thr His Thr
1 5 10

<210> 29
<211> 13
<212> PRT
<213> Peptococcus niger

<400> 29
Asp Pro Ile Asp Glu Ala Gly Glu Val Pro Ser Glu Thr
1 5 10

<210> 30
<211> 25
<212> PRT
<213> Homo sapiens

<400> 30
Asp Asp Asp Gly Ile Gly Tyr Val Glu Asp Gly Arg Glu Ile Phe Asp
1 5 10 15

Asp Asp Leu Glu Asp Asp Ala Leu Asp
20 25

<210> 31
<211> 59
<212> PRT
<213> Vargula sp.

<400> 31
Tyr Trp Asn Thr Trp Asp Val Lys Val Ser Leu Arg Asp Val Glu Ser
1 5 10 15

Tyr Thr Glu Val Glu Lys Val Thr Ile Arg Lys Gln Ser Thr Val Val
20 25 30

Asp Leu Ile Val Asp Gly Lys Gln Val Lys Val Gly Gly Val Asp Val
35 40 45

Ser Ile Pro Tyr Ser Ser Glu Asn Thr Ser Ile
50 55

<210> 32
<211> 62
<212> PRT
<213> Renilla sp.

<400> 32
Ala Ile Lys Ile Ala Lys Leu Ser Ala Glu Lys Ala Glu Glu Thr Arg
1 5 10 15

Gly Phe Leu Arg Val Ala Asp Gln Leu Gly Leu Ala Pro Gly Val Arg
20 25 30

Ile Ser Val Glu Glu Ala Ala Val Asn Ala Thr Asp Ser Leu Leu Lys
35 40 45

Met Lys Ala Glu Glu Lys Ala Met Ala Val Ile Gln Ser Leu
50 55 60

<210> 33
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Illustrative
P-loop binding motif

<400> 33
Ala Ala Ala Ala Gly Lys Thr
1 5

<210> 34
<211> 4
<212> PRT
<213> Photinus pyralis

<400> 34
His His Gly Phe
1

<210> 35
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic illustrative peptide

<400> 35
Met Leu Ser Arg Leu Ser Leu Arg Leu Leu Ser Arg Tyr Leu Leu
1 5 10 15

<210> 36
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic illustrative peptide

<400> 36
Lys Lys Ser Ala Leu Leu Ala Leu Met Tyr Val Cys Pro Gly Lys Ala
1 5 10 15

Asp Lys Glu

<210> 37
<211> 16
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic illustrative peptide

<400> 37
Met Leu Leu Pro Val Pro Leu Leu Leu Gly Leu Leu Gly Leu Ala Ala
1 5 10 15

<210> 38
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic illustrative peptide

<400> 38
Lys Asp Glu Leu
1

<210> 39
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic illustrative peptide

<400> 39
His Asp Glu Leu
1

<210> 40
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic illustrative peptide

<400> 40
Lys Glu Glu Leu
1

<210> 41
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic illustrative peptide

<400> 41
Pro Lys Lys Lys Arg Lys Val
1 5

<210> 42
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Illustrative
N-terminal acylation motif from Tyrosine kinase

<400> 42
Met Gly Cys Val Cys Ser Ser Asn Pro Asp
1 5 10